Grassmannian as a metric space

Subhajit Mishra
Department of Physics, IISER Kolkata

Under the supervision of
Dr. Somnath Basu
Department of Mathematics and Statistics, IISER Kolkata

November 26, 2016

1 Introduction

Definition 1.1. (Metric Space) Let X be a set. A function $d : X \times X \to \mathbb{R}$ is called *metric* if the following holds:

(i) $d(x, y) \geq 0 \quad \forall x, y \in X$ and $d(x, y) = 0 \iff x = y$;

(ii) $d(x, y) = d(y, x)$ (symmetric);

(iii) $d(x, y) \leq d(x, z) + d(z, y) \quad \forall x, y, z \in X$ (triangle inequality).

The set X equipped with a metric is called *metric space*, denoted by (X, d).

Definition 1.2. (Grassmannian of k-planes) Grassmannian of k-planes is a set of all k dimensional subspaces of $\mathbb{R}^n$. We denote it by $G_k(\mathbb{R}^n)$. From now on I will call any such as Grassmannian.

We define $d : G_k(\mathbb{R}^n) \times G_k(\mathbb{R}^n) \to \mathbb{R}^n$ such that

$$d(L, L') = \sup_{x \in L \cap S^{n-1}} d(x, L') \tag{1}$$

where $d(x, L') = \inf_{y \in L'} d_E(x, y)$, with $d_E(x, y)$ as the usual euclidean metric in $\mathbb{R}^n$.

We will show that $(G_k(\mathbb{R}^n), d)$ is a metric space and the induced topology has properties - compactness and path-connectedness.

2 Proof of $(G_k(\mathbb{R}^n), d)$ as a metric space

In this section we will show that $G_k(\mathbb{R}^n)$ is a metric space equipped the function d defined in Eqn (1). First, we show that d is a metric.

Claim. d is a metric

Proof. (i) We begin by showing that the first property of metric holds.

First we prove a result, which will be used later.

Result 2.1. Let V and W be two k dimensional subspaces of a vector space and $V \subseteq W$. Then $V = W$.

Proof. Let $\dim V = \dim W = k$ and $B_v = v_1, v_2, ..., v_k$ be a basis of V. As $V \subseteq W$, B_v is a list in W. Moreover, it is a linearly independent list in W (since, it is a basis of V).
Now, we can make B_v a basis of W by adding some vectors of W in B_v. But as $\dim W = k$, any basis of W must have length k. So, we can conclude B_v is also a basis of W. This implies, $\forall w \in W$, $w = \sum_{i=1}^{k} a_i v_i \in V$ (as B_v is basis of V). So, $W \subseteq V$. Hence, $V = W$. ☐

Now,

$$\begin{aligned} & d_E(x,y) \geq 0; \quad \text{(as it is the usual metric in } \mathbb{R}^n) \\ \implies & \inf_{y \in L'} d_E(x,y) \geq 0 \\ \implies & \sup_{x \in L \cap S^{n-1}} \inf_{y \in L'} d_E(x,y) \geq 0 \\ \implies & d(L,L') \geq 0 \end{aligned}$$

Since L and L' are are arbitrary, it is true for all L and L' in $G_k(\mathbb{R}^n)$.
Let $d(L,L') = 0$. This implies:

$$\begin{aligned} & \sup_{x \in L \cap S^{n-1}} d(x,L') = 0 \\ \implies & d(x,L') = 0, \quad \forall x \in L \cap S^{n-1} \\ \implies & \inf_{y \in L'} d_E(x,y) = 0 \\ \implies & d_E(x,y_0) = 0, \quad \text{(for some } y_0 \in L'\text{, by } \textbf{Claim 2.4}\text{, see next page)} \\ \implies & x = y_0 \\ \implies & x \in L' \\ \implies & L \subseteq L'. \end{aligned}$$

By **Result 2.1**, we get $L = L'$. So, $d(L,L') = 0 \implies L = L'$.
Now, consider $d(L,L) = \sup_{x \in L \cap S^{n-1}} \inf_{y \in L} d_E(x,y)$. Fix $x_0 \in L \cap S^{n-1}$. So,

$$\begin{aligned} & d_E(x_0,x_0) = 0; \\ \implies & \inf_{y \in L} d_E(x_0,y) = 0; \quad \text{(as } d_E(x,y) \geq 0) \\ \implies & \sup_{x_0 \in L \cap S^{n-1}} \inf_{y \in L} d_E(x_0,y) = 0; \\ \implies & d(L,L) = 0. \end{aligned}$$

So, $d(L,L') \geq 0$ and $d(L,L') = 0 \iff L = L'$.

(ii) To show that the function d is *symmetric*, first we show that the supremum is actually achieved i.e., $\sup_{x \in L \cap S^{n-1}} d(x,L') = d(x_0,L')$, for some $x_0 \in L \cap S^{n-1}$.

Lemma 2.2. The function $d : \mathbb{R}^n \to \mathbb{R}$ such that $d(x,L') = \inf_{y \in L'} d_E(x,y)$ is continuous in $\mathbb{R}^n$.

Proof. Let $\{x_n\}_{n \in \mathbb{N}}$ be a sequence in $\mathbb{R}^n$, which converges to $x_0 \in \mathbb{R}^n$ i.e., $\forall \epsilon > 0, \exists N \in \mathbb{N}$ such that $\forall n \geq N$, $d_E(x_n,x_0) < \epsilon$.
We fix n. $d(x_n,L') = \inf_{y \in L'} d_E(x_n,y)$ implies: given $\epsilon > 0$, $\exists y \in L'$ such that $d(x_n,L') \leq d_E(x_n,y) \leq d(x_n,L') + \epsilon$. Now,

$$\begin{aligned} d(x_0,L') \leq d_E(x_0,y) & \leq d_E(x_0,x_n) + d_E(x_n,y), \quad \text{(triangle inequality)} \\ & \leq d_E(x_0,x_n) + d(x_n,L') + \epsilon \\ & < d(x_n,L') + 2\epsilon, \quad (\forall n \geq N) \\ & \implies d(x_0,L') - d(x_n,L') < 2\epsilon, \quad (\forall n \geq N \text{ and } \forall \epsilon > 0) \end{aligned}$$

By interchanging x_0 and x_n we will get $d(x_n, L') - d(x_0, L') < 2\epsilon$, $\forall n \geq N$ and $\forall \epsilon > 0$. So, $|d(x_n, L') - d(x_0, L')| < 2\epsilon$. Hence $d(x, L')$ is continuous at x_0. Since it is continuous for any x_0 in $\mathbb{R}^n$, it is continuous in $\mathbb{R}^n$. □

Proposition 2.3. Let $g : X \to \mathbb{R}$ be a continuous function and X be compact. Then $\sup_{x \in X} g(x) \in g(X)$.

Proof. Since g is continuous and X is compact, $g(X) \subseteq \mathbb{R}$ is compact. So, $g(X)$ is closed in bounded (Heine-Borel Theorem). As $g(X)$ is bounded, $\sup g$ exists, and as $g(X)$ is closed, $\sup g \in g(X)$. □

Now, $L \cap S^{n-1} = S^{k-1}$ is compact. We restrict the function $d(x, L')$ to $L \cap S^{n-1}$. So, $d(x, L')$ is continuous in $L \cap S^{n-1}$ (subset of $\mathbb{R}^n$). By **Proposition 2.3**, supremum is achieved. So, $d(L, L') = \sup_{x \in L \cap S^{n-1}} d(x, L') = d(x_0, L')$, for some $x_0 \in L \cap S^{n-1}$.

Now, $d(x_0, L') = \inf_{y \in L'} d_E(x_0, y)$.

Claim 2.4. $\inf_{y \in L'} d(x_0, y) = d_E(x_0, y_0)$, for some $y_0 \in L'$, and y_0 is unique.

Proof. Let y_0 be the orthogonal projection of x_0 into L'. So, $\langle x_0 - y_0, y_0 \rangle = 0$. Moreover, as $x_0 - y_0$ is perpendicular to the plane L', we get $\langle x - y_0, v \rangle = 0, \forall v \in L'$.

Let $\{v_i\}_{i=1}^k$ be an orthonormal basis of L'. Since $y_0 \in L'$, $y_0 = \sum_{i=1}^k a_i v_i$, $a_i \in \mathbb{R}$. We extend this list to a basis $\{v_1, ..., v_k, u_{k+1}, ..., u_n\}$ of $\mathbb{R}^n$. By Gram-Schimdt orthogonalization, we get $\{v_i\}_{i=1}^n$, an orthonormal basis of $\mathbb{R}^n$.

Now, $x_0 \in \mathbb{R}^n \implies x_0 = \sum_{i=1}^n b_i v_i$, where $b_j \in \mathbb{R}$.

We take $v = v_j$, for $j = 1, 2, ..., k$.

$$\begin{aligned}
&\langle x - y_0, v_j \rangle = 0 \\
\implies &\left\langle \sum_{i=1}^n b_i v_i - \sum_{i=1}^k a_i v_i, v_j \right\rangle = 0 \\
\implies &\left\langle \sum_{i=1}^k (b_i - a_i) v_i, v_j \right\rangle = 0, \quad (\langle v_i, v_j \rangle = 0, \text{ for } i \neq j) \\
\implies &b_j = a_j \quad (\text{for } j = 1, 2, ..., k)
\end{aligned}$$

So, $y_0 = \sum_{i=1}^k b_i v_i$.

Let $y_0' \in L'$ and $y_0' \neq y_0$. So, $y_0' = \sum_{i=1}^k c_i v_i, \quad c_i \in \mathbb{R}$. Now,

$$\begin{aligned}
d_E(x_0, y_0') &= \sqrt{\sum_{i=1}^k (b_i - c_i)^2 + \sum_{i=k+1}^n b_i^2} \\
\implies (d_E(x_0, y_0'))^2 &= \sum_{i=1}^k (b_i - c_i)^2 + (d_E(x_0, y_0))^2 \\
\implies d_E(x_0, y_0') &> d_E(x_0, y_0)
\end{aligned}$$

Since y_0' is arbitrary, we can say that $\inf_{y \in L'} d(x_0, y) = d_E(x_0, y_0)$.

To show that y_0 is unique, let us assume $\exists \tilde{y} \in L'$ such that $\langle x_0 - \tilde{y}, \tilde{y} \rangle = 0$. Now, $x_0 = y_0 + y_0^\perp = \tilde{y} + \tilde{y}^\perp$. This gives, $y_0 - \tilde{y} = \tilde{y}^\perp - y_0^\perp$. Since $y_0 - \tilde{y} \in L'$ and $\tilde{y}^\perp - y_0^\perp \in L'^\perp$ and both are equal, $y_0 - \tilde{y} = 0 \implies y_0 = \tilde{y}$. Hence, y_0 is unique. □

So, $d(L, L') = d(x_0, L') = d_E(x_0, y_{x_0})$, where y_{x_0} is the orthogonal projection of x_0 into L'. Now,

$$\begin{aligned} d(L, L') &= d_E(x_0, y_{x_0}) \\ &= d_E(y_{x_0}, x_0), \quad \text{(symmetry)} \\ &= d_E(\hat{y}_0, x_{\hat{y}_0}) \quad \text{(using properties of congruent triangles)} \\ &= d(\hat{y}_0, L), \quad \text{(since, infimum is achieved)} \\ &\leq \sup_{y \in L' \cap S^{n-1}} d(y, L) = d(L', L). \end{aligned}$$

where $\hat{y}_0$ is the unit vector in the direction of y_{x_0} and $x_{\hat{y}_0}$ is the orthogonal projection of $\hat{y}_0$ on L.

Now, interchanging L' and L we get $d(L', L) \leq d(L, L')$. So, $d(L, L') = d(L', L)$; (symmetric).

(iii) Triangle inequality:

Let L, L', $L''\in G_k(\mathbb{R}^n)$. We need to show that $d(L, L'') \leq d(L, L') + d(L', L'')$.

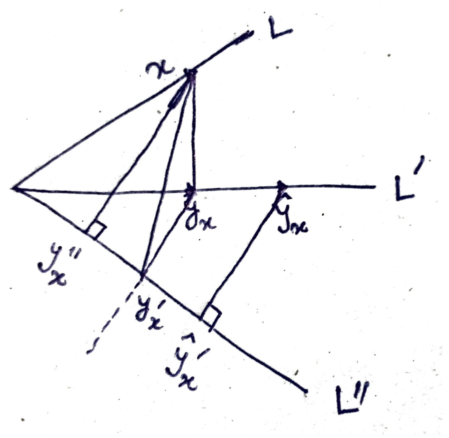

Figure 1: for $\|y_x\| \neq 0$

Take $x \in L \cap S^{n-1}$ such that $d(L, L'') = d_E(x, y''_x)$, where $y''_x = P_{L''}(x)$ is the orthogonal projection of x on L'' (see Figure 1, as supremum and infimum are achieved). Using the Pythagoras Theorem we get,

$$\begin{aligned} &\|x\|^2 = \|y_x\|^2 + \|x - y_x\|^2 \\ \implies &\|y_x\|^2 + \|x - y_x\|^2 = 1 \quad \text{(since } x \text{ is a unit vector)} \\ \implies &\|y_x\| \leq 1; \quad \text{(as } \|x - y_x\|^2 \geq 0\text{)} \end{aligned}$$

Now, $\hat{y}_x = \lambda y_x$, where $\lambda = \frac{1}{\|y_x\|} \geq 1$ ($\|y_x\| \neq 0$). So,

$$\begin{aligned} \|\hat{y}_x - P_{L'}(\hat{y}_x)\| &= |\lambda| \|y_x - P_{L'}(y_x)\| \quad \text{(as projection map } P \text{ is linear)} \\ &\geq \|y_x - P_{L'}(y_x)\| \end{aligned}$$

The distance between $\hat{y}_x$ and $\hat{y}'_x = P_{L'}(\hat{y}_x)$ is always greater than the distance between y_x and y'_x. Now, translate the vector $\hat{y}_x - P_{L'}$ to y_x so that the line starting from y_x intersects L'' at y'_x.

Since $d_E(x, y''_x)$ is the shortest distance between L and L'', we have

$$\begin{aligned} d_E(x, y''_x) &\leq d_E(x, y'_x) \\ &\leq d_E(x, y_x) + d_E(y_x, y'_x); \quad \text{(triangle inequality)} \\ &\leq d_E(x, y_x) + d_E(\hat{y}_x, \hat{y}'_x); \quad \text{(as } d_E(y_x, y'_x) \leq d_E(\hat{y}_x, \hat{y}'_x)\text{)} \end{aligned}$$

Since we have started with the x for which the supremum is achieved, from the above relation we have $d(L, L'') \leq d(L, L') + d(L', L'')$.

For the case, where $\|y_x\| = 0$ i.e., two of the planes are orthogonal to each other (say L and L', see Figure 2); we have $d(L, L'') \leq 1 = d(L, L')$ (since, x is a unit vector). So,

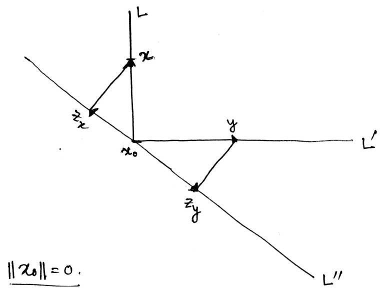

Figure 2:

$d(L, L'') \leq d(L, L') + d(L', L'')$ (as $d(L', L'') \geq 0$).

We have shown that the function d satisfies all the properties of metric. Hence, d is a metric and $(G_k(\mathbb{R}^n), d)$ is a metric space. □

The metric, d is uniformly bounded by 1. Since, the distance between any two k-planes is given by the distance between a unit vector and the orthogonal projection of the unit vector. The distance is 1 when there are two k-planes perpendicular to each other.

3 Some topological properties of Grassmannian

The induced topology on Grassmannian has some nice topological properties - compactness and path-connectedness. Hausdorff property of Grassmannian is evident (as it is a metric space). We check the two other topological properties, mentioned here.

Proposition 3.1. $G_k(\mathbb{R}^n)$ is compact.

Proof. Let $f : O(n) \to G_k(\mathbb{R}^n)$ be a function defined as $f(A) = span(A_1, A_2, ..., A_k)$, where A_i is a coloumn vector of A. If we can show that f is *onto* and *continuous*, then we are done.

Claim 3.2. *(f is onto) $\forall L \in G_k(\mathbb{R}^n)$, $\exists A \in O(n)$ such that $f(A) = L$.*

Proof. Let $B_L = \{v_i\}_{i=1}^k$ be an orthonormal basis of L. Since B_L is a linearly independent list, we extend it to $\{v_1, v_2, ..., v_k, u_{k+1}, ..., u_n\}$, a basis of $\mathbb{R}^n$. Using Gram-Schimdt orthogonalization we get an orthonormal basis $\{v_i\}_{i=1}^n$ of $\mathbb{R}^n$.

So taking $A = [v_1 \ v_2 \ ... \ v_n]$, we have $f(A) = span(v_1, v_2, .., v_k) = L$. As v_i's are orthonormal, $A \in O(n)$. Since L is arbitrary, f is *onto.* □

Claim 3.3. ($f : O(n) \to G_k(\mathbb{R}^n)$ is continuous) $\forall \epsilon > 0$, $\exists \delta > 0$ such that $\forall A \in O(n)$, $\forall B \in O(n)$ if $d_E(A, B) < \delta$ then $d(f(A), f(B) < \epsilon$.

(Note). We have taken the metric in $O(n)$ to be euclidean because we can identify an element of $O(n)$ as a vector of $\mathbb{R}^{n^2}$.

Proof. Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be elements of $O(n)$. Now,

$$\begin{aligned}
(d_E(A,B))^2 &= \sum_{(i,j)=1}^{n} (a_{ij} - b_{ij})^2 < \delta^2 \\
&\implies \sum_{i=1}^{n}(a_{i1} - b_{i1})^2 + ... + \sum_{i=1}^{n}(a_{in} - b_{in})^2 < \delta^2 \\
&\implies \sum_{i=1}^{n}(a_{ij} - b_{ij})^2 < \delta^2; \text{ since } (a_{ij} - b_{ij})^2 \geq 0, \text{ for each } j = 1, 2, ..., n \\
&\implies d_E(A_j, B_j) < \delta, \quad (\text{for each } j = 1, 2, ..., n)
\end{aligned}$$

Since supremum and infimum both are achieved, we have $d(f(A), f(B)) = \sup_{x \in f(A) \cap S^{n-1}} d(x, f(B))$ $= d(x, f(B)) = d_E(x, P_{f(B)}(x))$, for some $x \in f(A) \cap S^{n-1}$, where P is the orthogonal projection operator. Now, $x = \sum_{i=1}^{k} \lambda_i A_i$, with $\sum_{i=1}^{n} \lambda_i^2 = 1$. So,

$$\begin{aligned}
d(f(A), f(B)) &= d_E(x, P_{f(B)}(x)) \\
&= d_E\left(\sum_{i=1}^{k} \lambda_i A_i, P_{f(B)}\left(\sum_{i=1}^{k} \lambda_i A_i\right)\right) \\
&= d_E\left(\sum_{i=1}^{k} \lambda_i A_i, \sum_{i=1}^{k} \lambda_i P_{f(B)}(A_i)\right), \quad (\text{since, } P \text{ is linear}) \\
&\leq \sum_{i=1}^{k} d_E\Big(\lambda_i A_i, \lambda_i P_{f(B)}(A_i)\Big), \quad (\text{using triangle inequality}) \\
&= \sum_{i=1}^{k} |\lambda_i| d_E\Big(A_i, P_{f(B)}(A_i)\Big), \quad (\text{as } d_E \text{ is the euclidean norm}) \\
&\leq \sum_{i=1}^{k} |\lambda_i| d_E(A_i, B_i); \quad \Big(\text{as } d_E\Big(A_i, P_{f(B)}(A_i)\Big) \text{ is the infimum }\Big) \\
&< \sqrt{k}\delta \quad (\text{as } d_E(A_i, B_i) < \delta \text{ and using Cauchy-Schwarz inequality, } \sum_{i=1}^{k} |\lambda_i| \leq \sqrt{k}) \\
&= \epsilon \quad (\text{taking } \delta = \frac{\epsilon}{\sqrt{k}})
\end{aligned}$$

Since f is continuous for all A, B in $O(n)$, it is continuous in $O(n)$. □

As, $f : O(n) \to G_k(\mathbb{R}^n)$ is onto and continuous, and $O(n)$ is compact, we have $G_k(\mathbb{R}^n)$ is compact. □

Proposition 3.4. $G_k(\mathbb{R}^n)$ is path-connected.

Proof. Let $f : SO(n) \to G_k(\mathbb{R}^n)$ is a function defined as $f(A) = span(A_1, A_2, ..., A_n)$.
As $SO(n) \subset O(n)$ and f is continuous on $O(n)$ (shown in the previous section), f is continuous on $SO(n)$. It is sufficient to show that f is *onto*.

Going by the same logic as in the proof of **Claim 2**, we get an orthogonal matrix $A = [v_1\ v_2\ ...\ v_n]$. If $\det(A) = 1$, then $A \in SO(n)$ and we are done.

If $\det(A) = -1$, we do a linear transformation, to get A' such that

$$A' = [v_1\ v_2\ ...\ v_n] \begin{bmatrix} -1 & 0 & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & 0 & ... & 1 \end{bmatrix} = [-v_1\ v_2\ ...\ v_n]$$

Now, $\det(A') = 1$. So, $A' \in SO(n)$. Hence, f is onto.

We know that continuous function preserves path-connectedness. Since $SO(n)$ is path-connected and $f : SO(n) \to G_k(\mathbb{R}^n)$ is *continuous* and *onto*, $G_k(\mathbb{R}^n)$ is path-connected. $\square$